

doi: 10.7690/bgzdh.2020.10.020

一种改进的灰色 BP 神经网络预测模型

柴志君, 欧阳中辉, 岳炯

(海军航空大学岸防兵学院, 山东 烟台 264001)

摘要: 为提高灰色 GM(1,1)模型的预测精度和 BP 神经网络的映射能力, 提出一种改进的灰色 BP 神经网络预测模型。通过分别对灰色系统理论和 BP 神经网络 2 种模型进行改进, 再串联组合成新的预测模型, 并结合实测数据进行仿真试验。结果表明: 改进后的模型能兼具二者优点, 预测精度有较大提高, 相对误差减小, 运算速率更快。

关键词: 灰色预测模型; BP 神经网络; 预测精度; 性能预测

中图分类号: TP302 文献标志码: A

An Improved Prediction Model of Grey BP Neural Network

Chai Zhijun, Ouyang Zhonghui, Yue Jiong

(College of Shore Defense, Navy Aviation University, Yantai 264001, China)

Abstract: In order to improve the prediction accuracy of grey GM(1,1) model and the mapping ability of BP neural network, an improved prediction model of grey BP neural network is proposed. Two models, including grey system theory and BP neural network, were improved respectively, and then a new prediction model was combined in series. Carry out the simulation by test data. The results show that the improved model has the advantages of both methods, the prediction accuracy is highly improved, the relative error is reduced, and the operation speed is faster.

Keywords: grey prediction model; BP neural network; prediction accuracy; performance prediction

0 引言

灰色神经网络的原理是将灰色系统理论和神经网络有机结合, 主要应用于解决数据量较少的不确定性问题^[1-3]。灰色系统理论是我国著名学者邓聚龙于 1982 年提出的一个采用数据生成, 并从生成中寻找数学规律的边缘学科^[4-5]。灰色系统理论的精神是通过对小数据量样本进行灰色生成, 对小数据、不确定系统进行处理加工, 从中提取有价值的信息知识, 发掘其中规律^[6]。神经网络是一种人工智能新兴起的研究方法, 具有大规模并行、分布式存储和处理、自组织、自适应和自学能力, 特别是处理需要同时考虑许多因素和条件的、不精确和模糊的信息处理问题^[7]。2 种方法针对不同预测对象各有优点, 灰色预测模型通常应用于规律趋势较强、波动不明显的预测问题, 神经网络主要应用于规律无序、波动相差较大的时间序列^[8-9]。电子器件广泛应用于各科学技术领域, 特别是航空航天系统对电子器件的高性能、高可靠、长性能, 甚至“零故障”要求更加严格^[10]。电子器件在长期贮存过程中, 由于受到温度、湿度、机械振动以及电化学腐蚀等外界因素变化的影响, 内部材料性能也随之发生物理

和化学变化, 对应的性能参数也将改变, 从而造成使用障碍, 甚至失效^[11]。目前, 灰色神经网络在常规电子器件性能分析及性能预测方面已有较广泛的应用, 例如, 陈华平等^[12]对电子产品性能评估方法现状、数学模型和加速实验的方法进行了详细分析, 并介绍了灰色预测和神经网络预测方法的研究现状和发展前景。韩建立等^[13]结合灰色系统理论和 BP 神经网络建立了灰色神经网络组合预测模型, 对某型空舰导弹电子设备失效时间进行预测, 结果显示: 预测精度高且实用性强, 效果明显优于传统灰色预测模型。

1 灰色 GM(1,1)模型

1.1 传统灰色 GM(1,1)模型

灰色 GM(1,1)预测模型表达式为:

$$X^{(0)}(i) + aZ^{(1)}(i) = b \quad (1)$$

式中: $X^{(0)}(i)$ 为原始序列; a 为发展系数; b 为灰作用量; $Z^{(1)}(i)$ 为白化背景序列, 背景值 $z^{(0)} = (z^{(0)}(1), z^{(0)}(2), \dots, z^{(0)}(n))$, 其中

$$z^{(1)}(k) = \alpha x^{(1)}(k-1) + (1-\alpha)x^{(1)}(k), k = 2, 3, \dots, n \quad (2)$$

式中 α 一般取 0.5。设 $X^{(0)}$ 是非负序列, $X^{(1)}$ 是 $X^{(0)}$

收稿日期: 2020-05-16; 修回日期: 2020-06-21

作者简介: 柴志君(1989—), 男, 山东人, 硕士, 从事军用仿真技术研究。E-mail: 275131218@qq.com。

的一次累加序列， $Z^{(1)}$ 是 $X^{(1)}$ 紧邻均值生成序列，则

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (3)$$

为灰色微分方程

$$x^{(0)}(t) + az^{(1)}(t) = b \quad (4)$$

的白化方程，也称影子方程。

采用最小二乘法，则离散型灰色微分方程 $x^{(0)}(k) + az^{(1)}(k) = b$ 的参数列满足：

$$[a, b]^T = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{Y}。 \quad (5)$$

式中：

$$\mathbf{B} = \begin{bmatrix} -\frac{1}{2}(X^{(1)}(1) + X^{(1)}(2)) & 1 \\ -\frac{1}{2}(X^{(1)}(2) + X^{(1)}(3)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(X^{(1)}(n-1) + X^{(1)}(n)) & 1 \end{bmatrix} = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -Z^{(1)}(n) & 1 \end{bmatrix};$$

$$\mathbf{Y} = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)]^T。$$

求方程的解：

$$\hat{x}^{(1)}(k+1) = (x^{(1)}(1) - b/a)e^{-ak} + b/a, \quad k=1, 2, \dots, n。 \quad (6)$$

对 $\hat{x}^{(1)}(k+1)$ 作一次累减，得到最终对应于原始数据的预测值 $\hat{x}^{(0)}(k+1)$ ，则

$$\begin{aligned} \hat{x}^{(0)}(k+1) &= \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) = \\ &= (1 - e^{-a})[x^{(1)}(1) - b/a]e^{-ak}, \\ &\quad k=1, 2, \dots, n。 \end{aligned} \quad (7)$$

1.2 无偏灰色 GM(1,1) 模型

在 GM(1,1) 指数序列建模过程中参数不可避免存在偏差。为了消除 GM(1,1) 模型在原始数据增长过程中可能失效的问题，采取对参数进行修正的方法，生成无偏灰色 GM(1,1) 模型，记作 WPGM(1,1)。WPGM(1,1) 同时也简化了计算过程中需要累减处理的环节，提高了计算速度。相对比传统 GM(1,1)，WPGM(1,1) 模型引入模型参数 A 和 B，其中，

$$A = 2b/(2+a), \quad B = \ln[(2-a)/(2+a)]。 \quad (8)$$

由此，可得到对应于原始数据序列的预测值的模型：

$$\hat{x}^{(0)}(k+1) = Ae^{Bk}, \quad k=1, 2, \dots, n。 \quad (9)$$

其中 $\hat{x}^{(0)}(1)=x^{(0)}(1)$ 。

1.3 改进无偏灰色 GM(1,1) 模型

对无偏灰色模型改进的常见传统方法有幂函数

法和新陈代谢法 2 种。幂函数法的原理是对序列数据分别求其二次方根，建立新的数据序列后再建模，能有效提高数列的光滑度和解决偏差问题。新陈代谢法的原理是在对原始样本补充新数据的同时，对原始样本中陈旧数据进行过滤，目的是使得改进后的数据序列的变化能够更好地适应运算模型，进一步提高运算精度。笔者将幂函数法和新陈代谢法相结合，对无偏灰色 GM(1,1) 模型进行改进，具体步骤如下：

1) $X^{(0)}(k)$ 为原始数据样本序列，求 $X^{(0)}(k)$ 中各数据二次方根，则得到变换后的数据序列 $X_m^{(0)}(k)$ ；

2) 对 $X_m^{(0)}(k)$ 进行一次累加后生成新的数据序列 $X_m^{(1)}(k)$ ， $X_m^{(1)}(k)$ 是 $X_m^{(0)}(k)$ 的 1-AGO 序列；

3) 按照式(8)求得改进后无偏灰色 GM(1,1) 模型的预测值 $x_m^{(0)}(k+1)$ ，再进行二次乘方运算，则 $[x_m^{(0)}(k+1)]^2$ 为实际样本模型的预测值；

4) 将 $[x_m^{(0)}(k+1)]^2$ 放入原始数据样本序列 $X^{(0)}$ 中，同时利用新陈代谢剔除 $X^{(0)}$ 中第 1 个数据 $x^{(0)}(1)$ ，则得到新的数据序列 $\hat{X}_m^*(1)$ ；

5) 按照上面步骤进行数据替换，则改进算法生成新数据序列 $\hat{X}_m^*(k)$ ，可得到通过 n 次替换的预测值为 $\hat{x}_m^*(n)$ 。

1.4 模型精度检验

模型精度反映模型预测的精确度，在初步选定模型后，要对各模型的精度进行检验，来判断其是否合理，经过检验合格的模型才可作预测。模型方法有关联度检测、残差检测和后验误差检验模型（包括均方法检测和小误差概率检测）。对应的检测指标和相应的模拟精度等级见表 1^[14]。

表 1 精度检验等级

精度等级	关联度 ε	相对误差 α	均方差比值 C	小误差概率 ρ
1	0.90	0.01	0.35	0.95
2	0.80	0.05	0.50	0.80
3	0.70	0.10	0.65	0.70
4	0.60	0.20	0.80	0.60

2 BP 神经网络

2.1 传统 BP 神经网络模型

BP (back propagation) 神经网络采用误差反向传播的方法来调整各层权值阈值，传统 BP 网络主要由 1 个输入层、1 个或多个隐含层和 1 个输出层

组成, 隐含层中的神经元采用 sigmoid 型传递函数, 输出层函数采用纯线性传递函数, 结构如图 1。

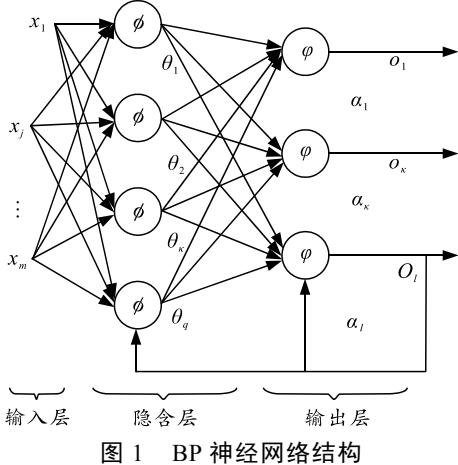


图 1 BP 神经网络结构

其推导过程如下:

1) 输入信号正向传播过程。

隐含层第 i 个节点的输入

$$\text{net}_i = \sum_{j=1}^m w_{ji} x_j + \theta_i。 \quad (10)$$

输出层函数为线性传递函数, 第 k 个节点的输入为

$$\text{net}_k = \sum_{i=1}^q w_{ik} y_i + a_k。 \quad (11)$$

其中 $y_i = \phi(\text{net}_i)$ 。

第 k 个节点的输出 O_k 为

$$o_k = \phi(\text{net}_k) = \phi\left(\sum_{i=1}^q w_{ik} \phi\left(\sum_{j=1}^m w_{ji} x_j + \theta_i\right) + a_k\right)。 \quad (12)$$

定义误差函数

$$E_k = \frac{1}{2} \sum_{k=1}^l (F_k - o_k)^2。 \quad (13)$$

其中: F_k 为第 k 个输出的真实值; E_k 为输出误差。预测值与真实值之间误差越小, 说明预测精度越高。

2) 反向传播过程。

隐含层权值公式为

$$w_{ji} = w_{ji}(t) + \eta \delta_j o_k + \alpha [w_{ji}(t) - w_{ji}(t-1)]。 \quad (14)$$

输出层权值公式为

$$w_{ik} = w_{ik}(t) + \eta \delta_i o_k + \alpha [w_{ik}(t) - w_{ik}(t-1)]。 \quad (15)$$

式中: η 为学习率, $\eta \in (0,1)$; α 为动量因子, $\alpha \in (0,1)$; t 为调节次数; δ 与偏差有关。

对于输出层:

$$\delta_i = y_i(1-y_i)(F_i - y_i)。 \quad (16)$$

对于隐含层:

$$\delta_j = x_j(1-x_j) \sum_{i=0}^q \delta_i w_{ji}。 \quad (17)$$

2.2 BP 神经网络参数的改进

与常规权值调整公式中常数项动量因子 α_0 不同, 改进后的动量因子是一个变量, 其取值受前面输出误差的影响, 大小随输出误差比值大小作调整。按照经验可知: $\alpha_0 \in (0,1)$, 误差调整总体呈下降趋势。

现令

$$\alpha(t) = \alpha_0 \cdot E(t-1) / E(t)。 \quad (18)$$

式(14)、式(15)调整为:

$$w_{ji} = w_{ji}(t) + \eta \delta_j o_k + [\alpha_0 \cdot E(t-1) / E(t)] [w_{ji}(t) - w_{ji}(t-1)]; \quad (19)$$

$$w_{ik} = w_{ik}(t) + \eta \delta_i o_k + [\alpha_0 \cdot E(t-1) / E(t)] [w_{ik}(t) - w_{ik}(t-1)]。 \quad (20)$$

同理可得阈值的调整公式, 在此不作说明。

此外, 学习速率 η 的取值对网络的训练性能也同样有一定影响, 通常为一常数: η 取值过小, 会延长训练时间, 训练收敛减慢; η 取值过大, 对网络稳定性降低起一定作用。综合考虑训练时间和稳定性 2 方面因素, 参照预先设定的误差参数, 提出对 η 取值进行调整, 当网络输出误差较上一层输出超出设定值, 则减小 η 值, 反之增加, 直至网络训练收敛达到预期。 η 取值如下:

$$\eta(t) = \begin{cases} 1.05\eta(t-1), & E(t) < E(t-1) \\ 0.7\eta(t-1), & E(t) > 1.04E(t-1) \\ \eta(t-1), & \text{其他} \end{cases}。 \quad (21)$$

通过自动调整 η 来实现始终以最大允许速率对网络进行训练。

3 改进型灰色 BP 神经网络

3.1 灰色 BP 神经网络组合形式

灰色 BP 神经网络是将灰色系统理论和 BP 神经网络结合在一起, 并对不确定的问题进行求解的模型^[15]。灰色理论模型与神经网络传统结合方式主要有以下几种^[16]:

1) 串联型: 对多个灰色预测模型预测的结果, 再通过神经网络系统进行预测;

2) 并联型: 先采用灰色预测模型和神经网络系统分别进行预测, 再对所有预测结果结合实际背景加权组合作为最终预测值;

3) 嵌入型：在神经网络系统的输入层添加 1 个灰化层，在输出层添加 1 个白化层。

笔者建立的改进型灰色 BP 神经网络模型采用串联式组合模型，由改进后无偏灰色 GM(1,1)模型和改进后的 3 层 BP 神经网络串联而成。

3.2 改进型灰色 BP 神经网络组合预测算法步骤

1) 输入原始数据序列 $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ ；

2) 用前面 n 个样本数据分别建立 WPGM(1,1)模型和改进后的 WPGM(1,1)模型，用 2 个模型对原始数据分别进行预测，得到预测数据序列 $\hat{X}^{(0)} = (\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(n))$ ，同时利用式(13)建立残差数据序列 $E = X^{(0)} - \hat{X}^{(0)} = (e(1), e(2), \dots, e(n))$ ；

3) 对 2 个模型分别进行精度检验，如精度在允许范围内则可，如精度超出允许范围，则模型不可作预测；

4) 将预测数据序列 $\hat{X}^{(0)}$ 作为输入样本，输入 BP 神经网络模型和改进后的 BP 神经网络模型中，用残差数据序列 E 作为期望输出，通过足量的样本输入和期望输出，对改进前后的 BP 神经网络进行训练；

5) 利用训练好的 BP 神经网络对残差数据进行预测，则可以得到修正后的残差数据序列 $\hat{E} = (\hat{e}(1), \hat{e}(2), \dots, \hat{e}(n))$ ，则灰色 BP 神经网络的最终预测值：

$$Y^{(0)} = \hat{X}^{(0)} + \hat{E} = (y^{(0)}(1), y^{(0)}(2), \dots, y^{(0)}(n))。$$

改进型灰色 BP 神经网络预测结构如图 2 所示。

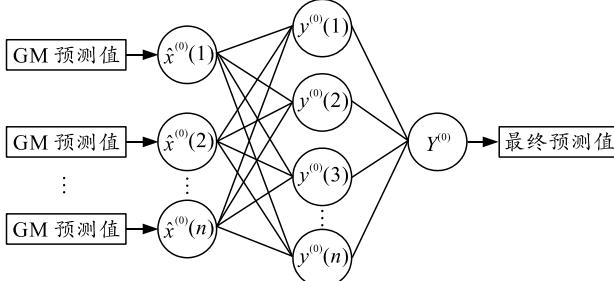


图 2 改进型灰色 BP 神经网络预测结构

4 实例应用

结合某型防空导弹雷达导引头发射机某电子器件性能的实测数据，对大量实测数据进行分区间处理，共分为 12 组实测数据，各区间数据经处理后其平均值分别是 1.951 1, 1.891 3, 1.790 7, 1.671 8,

1.634 4, 1.548 1, 1.450 9, 1.391 3, 1.339 4, 1.220 3, 1.113 6, 1.028 0。取前 10 个数据作为训练样本，经过 GM(1,1)、WPGM(1,1)和 GWPGM(1,1)建立模型，再将 10 个数据经处理后对应的残差作为期望输出样本。利用 10 个区间的原始数据分别对传统 BP 神经网络和改进后 BP 神经网络进行训练，经训练后的组合预测模型对后 2 个区间平均数据进行预测。

4.1 几种灰色 GM(1,1)预测模型

4.1.1 传统灰色 GM(1,1)模型预测

将训练样本代入式(1)—(5)，可得到灰色 GM(1,1)模型参数

$$[a, b]^T = [0.0258, 1.4276]^T。 \quad (22)$$

由式(6)和式(7)可得灰色 GM(1,1)模型预测值为：

$$x_{k+1}^{(0)} = \left(1 - e^{0.0258}\right) \left[1.9511 - \frac{1.4276}{0.0258}\right] e^{-0.0258k}。 \quad (23)$$

4.1.2 WPGM(1,1)模型预测

将式(28)所得模型参数 $[a, b]^T$ 代入式(8)，可得无偏灰色 GM(1,1)模型参数

$$A = 2.0139, B = -0.0563。 \quad (24)$$

由式(9)可得无偏灰色 GM(1,1)模型预测值为：

$$\hat{x}^{(0)}(k+1) = 2.0139e^{-0.0563k}, k = 1, 2, \dots, n。 \quad (25)$$

4.1.3 改进后的 GWPGM(1,1)预测模型

对训练样本各组数据求其二次方根，生成新的训练样本 $X_m^{(0)}(k)$ ，将新生成训练样本代入式(1)—(9)，按照 WPGM(1,1)预测模型步骤可得到预测值

$$x_m^{(0)}(k+1) = 1.3735e^{-0.0258k}。 \quad (26)$$

将 $[x_m^{(0)}(k+1)]^2 = 1.8865e^{-0.0516k}$ 代入原始数据序列 $X^{(0)}$ ，采用新陈代谢法经过一轮替换，建立新的数据序列 $X_m^*(1)$ ，循环上面步骤，则可得到改进后无偏灰色 GM(1,1)模型预测值。

得到 3 种预测模型样本数据的拟合数值见表 2。

表 2 不同模型拟合的结果

序号	实测数据平均值	GM	WPGM	GWPGM
1	1.951 1	1.951 1	1.951 1	1.951 1
2	1.891 3	1.906 6	1.799 5	1.886 4
3	1.790 7	1.802 3	1.701 1	1.791 4
4	1.671 8	1.703 7	1.608 0	1.701 2
5	1.634 4	1.610 5	1.520 0	1.615 6
6	1.548 1	1.522 4	1.436 8	1.534 2
7	1.450 9	1.439 2	1.358 2	1.457 0
8	1.391 3	1.360 4	1.283 9	1.383 6
9	1.339 4	1.286 0	1.213 7	1.314 0
10	1.220 3	1.215 7	1.147 2	1.247 8

参照表 1 对上述预测模型进行精度检验, 以判断 3 种模型的合理性。各项精度检验情况见表 3。

表 3 不同模型各项精度检验值

模型	关联度 ε	平均相对误差 $\alpha/\%$	均方差比值 C	小误差概率 ρ
GM	0.593 9	1.90	0.107 9	1.0
WPGM	0.519 9	5.02	0.144 0	1.0
GWPGM	0.733 8	2.09	0.124 9	1.0

由表可见: 3 种预测模型中传统灰色 GM(1,1)和无偏灰色 GM(1,1)关联度超出允许范围, 故不适用于此类预测; 改进后无偏灰色 GM(1,1)在 4 项指标上均符合精度检验, 故适用于此类预测。

4.2 灰色 BP 神经网络组合预测

将改进无偏灰色 GM(1,1)预测模型对原始数据产生的拟合值每 10 个数据分成一组, 作为 BP 神经

网络训练的输入值, 将下一数据拟合值与样本值的残差作为 BP 神经网络训练的输出值。BP 神经网络隐含层节点按照经验公式选取 15, 则 BP 神经网络的结构为 $10 \times 15 \times 1$ 。设网络训练次数为 2 000 次, 训练目标 0.000 1, w 的初始取值在 $(-0.3, +0.3)$ 内选取, $\eta=0.05$, $\theta_i(0)=0.9$, $a_k(0)=0.9$, 隐含层中的神经元采用 sigmoid 型传递函数, 输出层函数采用纯线性传递函数, 对传统 BP 神经网络和改进后 BP 神经网络进行训练。

4.3 预测结果比较

利用改进后无偏灰色 GM(1,1)和传统 BP 神经网络、改进后 BP 神经网络组合预测模型对目标数据进行预测, 共得到 3 组预测数据。预测数据与真值对应的相对误差见表 4。

表 4 3 种预测模型预测结果

序号	实测数据 平均值	GWPGM		GWPGM-BP		GWPGM-GBP	
		预测数据	相对误差/%	预测数据	相对误差/%	预测数据	相对误差/%
11	1.113 6	1.185 0	2.527	1.106 0	0.680	1.112 0	0.150
12	1.028 0	1.125 3	3.890	1.018 6	0.910	1.028 7	0.072

由表可见: 在每个数据区间中, 灰色 BP 神经网络组合预测相对误差明显小于改进后无偏灰色 GM(1,1)预测。针对 BP 神经网络环节的改进, 使得综合预测效果更佳, 预测值更接近实际值, 对该部件性能预测更符合发展走向。说明改进后的灰色 BP 神经网络对某型防空导弹发射机电子器件的性能预测具有更好的适用性。

5 结论

笔者对灰色 GM(1,1)预测模型进行改进, 分别利用传统灰色 GM(1,1)、无偏灰色 GM(1,1)和改进后的无偏灰色 GM(1,1), 结合实测数据对某型防空导弹雷达导引头发射机中某电子器件的性能进行预测。在传统 BP 神经网络的基础上通过调整动量因子和学习速率的取值, 对 BP 神经网络进行改进。将灰色 GM(1,1)预测模型和改进前后的 BP 神经网络结合成新的预测模型, 再对该电子器件进行性能预测。预测结果表明: 改进后的组合预测模型预测精度更高, 适用性更强。

参考文献:

- [1] 程序, 颜廷利, 詹花茂. GIS 中局部放电 UHF 信号传播特性研究[J]. 高压电器, 2014, 50(5): 30–35.
- [2] 彭仲秋. 瞬变电磁场[M]. 北京: 高等教育出版社, 1989: 99–121.
- [3] JUDD M D, FARISH O, HAMPTON B F. the Excitation

of UHF Signals by Partial Discharge in GIS[J]. IEEE Transaction on Dielectrics and Electrical Insulation, 1996, 35(126): 27–33.

- [4] LI Y, XU H J, ZHANG Y Z. The Grey Dynamic Model and ITS Application in the Grinding Temperature. The First Asia-Pacific and Second Japan-China International Conference on Progress of Cutting and Grinding, Shanghai, China, 1994: 108–115. Environment[J]. IEEE Transactions on Nuclear Science, 2007, 54(4): 982–988.
- [5] 刘思峰, 郭天榜, 党耀国. 灰色系统理论及其应用[M]. 北京: 北京科学出版社, 1999: 15–30.
- [6] 周志文. 灰色 BP 神经网络的研究[D]. 厦门: 厦门大学, 2009: 10–15.
- [7] 闻新, 陈镝, 乔羽. 国内神经网络故障诊断技术及其在航天器中的应用[J]. 沈阳航空航天大学学报, 2018, 35(3): 17–25.
- [8] LEIGHTON R R, CONRATH B C. The Autoregressive Backpropagation Algorithm[C]//Neural Networks, 1991. IJCNN-91-Seattle International Joint Conference on. IEEE, 1991: 369–377.
- [9] SHIMAKURA Y, FUJISAWA Y, MAEDA Y. Short-term Load Forecasting Using an Artificial Neural Network[C]//International Forum on Neural Networks to Power Systems, Annps. IEEE, 1993.
- [10] FRANCO F J, ZONG Y, AGAPITO J A. Degradation of Power Bipolar Operational Amplifiers in a Mixed Neutron and Gamma Environment[J]. IEEE Transactions on Nuclear Science, 2007, 54(4): 982–988.

(下转第 96 页)